

# Demultiplexing Report

Angela Crabtree

8/11/2021

## Distribution of Read Pairs in Demultiplexed Bins

Inputs for the python script, `demux.py`, were 4 FASTQ files (R1, R2, R3, R4), one tab-delimited text file with barcode id and sequence information, and a user-specified q-score cutoff. The cutoff determined which reads failed to have good-quality barcode sequences and were thus relegated to the “junk” bin. Note that this cutoff does NOT apply to the sequencing coverage of the reads themselves.

Here are the command-line arguments used:

```
## ./demux.py -1 /projects/bgmp/shared/2017_sequencing/1294_S1_L008_R1_001.fastq.gz -2 /projects/bgmp/s
```

If barcodes in FASTQ files did not match barcodes in the barcode info file, they were also binned in “junk”. If both forward and reverse barcodes for a read pair matched those in info file, but were not the same, read swapping occurred and read pairs were binned in “swap” bins. If barcodes matched, they were binned in the appropriate index bin. Counts were recorded and written to a csv output file.

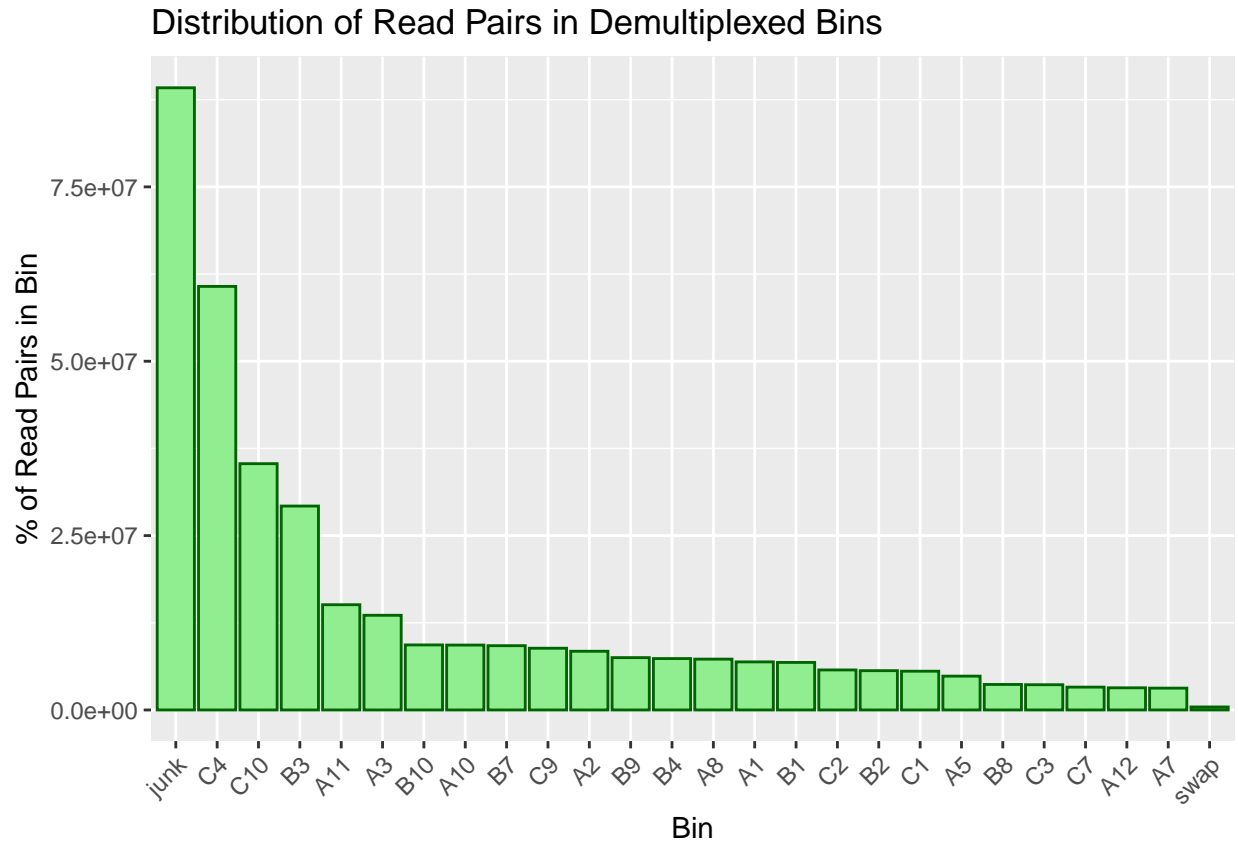
Here are the counts of read pairs binned:

bin	rpcount	rpp
junk	89210487	24.5591986
swap	432046	0.1189401
B1	6816735	1.8766129
A5	4852641	1.3359077
C1	5559436	1.5304848
B9	7505874	2.0663294
C9	8852290	2.4369909
C3	3620467	0.9966964
B3	29237911	8.0490499
C4	60731667	16.7191226
A11	15091465	4.1546044
C7	3277838	0.9023723
B2	5634659	1.5511933
A1	6893551	1.8977599
B7	9219872	2.5381844
A3	13578756	3.7381633
B4	7370239	2.0289897
A12	3176139	0.8743751
C10	35312354	9.7213135
A2	8419208	2.3177656
C2	5737265	1.5794402
A10	9307236	2.5622353

bin	rpcount	rpp
B8	3659488	1.0074386
A7	3129988	0.8616700
B10	9327234	2.5677406
A8	7291889	2.0074204

bin = index id or bin name; rpcount = read pair count; rpp = read pair percent of total

## Distribution of Read Pairs in Demultiplexed Bins



## Other Stats

Below are summary statistics:

## Percent of Junk Reads: 24.5592 %

## Percent of Swapped Reads: 0.1189401 %

##	bin	rpcount	rpp
##	Length:26	Min. : 432046	Min. : 0.1189
##	Class :character	1st Qu.: 5029340	1st Qu.: 1.3846
##	Mode :character	Median : 7331064	Median : 2.0182

##	Mean	:13971028	Mean	: 3.8462
##	3rd Qu.:	9322234	3rd Qu.:	2.5664
##	Max.	:89210487	Max.	:24.5592